

# Approximations and reducts with covering generalized rough sets

Eric C.C. Tsang<sup>a,\*</sup>, Chen Degang<sup>b</sup>, Daniel S. Yeung<sup>a</sup>

<sup>a</sup> *Department of Computing, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong*

<sup>b</sup> *Department of Mathematics and Physics, North China Electric Power University(Beijing), 102206, Beijing, PR China*

Received 15 September 2005; accepted 12 December 2006

---

## Abstract

The covering generalized rough sets are an improvement of traditional rough set model to deal with more complex practical problems which the traditional one cannot handle. It is well known that any generalization of traditional rough set theory should first have practical applied background and two important theoretical issues must be addressed. The first one is to present reasonable definitions of set approximations, and the second one is to develop reasonable algorithms for attributes reduct. The existing covering generalized rough sets, however, mainly pay attention to constructing approximation operators. The ideas of constructing lower approximations are similar but the ideas of constructing upper approximations are different and they all seem to be unreasonable. Furthermore, less effort has been put on the discussion of the applied background and the attributes reduct of covering generalized rough sets. In this paper we concentrate our discussion on the above two issues. We first discuss the applied background of covering generalized rough sets by proposing three kinds of datasets which the traditional rough sets cannot handle and improve the definition of upper approximation for covering generalized rough sets to make it more reasonable than the existing ones. Then we study the attributes reduct with covering generalized rough sets and present an algorithm by using discernibility matrix to compute all the attributes reducts with covering generalized rough sets. With these discussions we can set up a basic foundation of the covering generalized rough set theory and broaden its applications.

© 2008 Elsevier Ltd. All rights reserved.

**Keywords:** Rough sets; Covering generalized rough sets; Reduct; Discernibility matrix

---

## 1. Introduction

The concept of rough set was originally proposed by Pawlak [1] as a mathematical approach to handle imprecision, vagueness and uncertainty in data analysis. This theory has been demonstrated to have its usefulness and versatility in successfully solving a variety of problems [6–8]. The theory of rough sets deals with the approximation of an arbitrary subset of a universe by two definable or observable subsets called lower and upper approximations. By using the concepts of lower and upper approximations in rough set theory, knowledge hidden in information systems may be unraveled and expressed in the form of decision rules [2–5]. Another application of rough set theory is that of attribute reduct in databases. Given a dataset with discretized attribute values, it is possible to find a subset of the original attributes that contain the same information as the original one. The concept of attributes reduct can be

---

\* Corresponding author.

E-mail address: [csetsang@comp.polyu.edu.hk](mailto:csetsang@comp.polyu.edu.hk) (E.C.C. Tsang).

viewed as the strongest and the most important result in rough set theory to distinguish itself from other theories. However, as pointed out by some scholars, partition or equivalence relation, as the indiscernibility relation in the traditional rough set theory, is still restrictive for many applications, i.e. many practical data sets can not be dealt with by traditional rough sets. Some generalizations of rough sets could be found in the literature. One method is the relaxation of the equivalence relation. For example, Slowinski and Vanderpooten [9] proposed the substitution of the equivalence relation that models indiscernibility, with a relation with which only the reflexivity property is required. They called this type of relation a similarity relation, as they argued that transitivity and symmetry are not always essential. Another approach is the relaxation of the partition to a cover. In [10–12,14–17] the concept of a cover of a universe was presented to construct the upper and lower approximations of an arbitrary set. In [10] the authors mainly studied the structure of covers while in [11] the authors examined the relationship between the upper and lower approximation operators and some axioms satisfied by the traditional rough sets. However their definitions of upper approximation are not the same and they do not seem to be the most reasonable one. In [12] based on the definitions of upper and lower approximations in [10], the authors claimed that they have studied the reduct of covering generalized rough sets, but their reduct with covering generalized rough sets is just to reduce the “redundant” members in a cover and find the “smallest” cover that induces the same covering lower and upper approximations, i.e. they do not reduce a redundant cover from a family of covers. This is quite different from the purpose of reduct with traditional rough sets to delete dispensable attributes in a database. Other types of attributes reduct include attributes reduct with tolerance relation [3] for incomplete information system and with fuzzy rough sets for fuzzy information systems [18, 22], these reducts employ the idea of dependency function in traditional rough sets to define and compute reduct and the structure of reduct have not been studied in detail. Comparing with the study of reduct with traditional rough sets with respect to equivalence relation [13], less work on the reduct with covering generalized rough sets could be found. It is well known that the approximations of arbitrary sets and reduct of attributes are the most fundamental issues in rough set theory, and the development on these topics are obviously important to the covering generalized rough set theory. In many practical problems we always deal with covers instead of partitions which will be discussed in Section 3. Thus the research of the covering generalized rough sets is clearly of practical importance. In this paper we first present some suitable applied background for covering rough sets by proposing three kinds of data set that the traditional rough sets cannot handle and develop some methods to construct cover by data sets. We then conclude the existing research on set approximations in covering generalized rough sets and mainly discuss the definitions of upper approximation of the covering generalized rough sets. For a cover we define its induced cover and the intersection of a family of covers by using the induced cover. So the dispensable covers in a family of covers are certainly defined as the ones that without them the intersection of the left covers is invariant. We then define the discernibility matrix and develop an algorithm to compute all the reducts for a family of covers, according to the author’s knowledge, this is the first research work on this topic. Our research on the reduct with covering generalized rough sets is the natural generalization of the one in traditional rough sets and is quite different from the one in [12]. With the above discussion, we can develop the foundation of further research on covering generalized rough sets and its applications.

## 2. The fundamentals of Pawlak rough sets

In this section we recall the basic definitions of traditional rough set theory.

Let  $U$  denote a finite and nonempty set called the universe. Suppose  $R \subseteq U \times U$  is an equivalence relation on  $U$ , i.e.  $R$  is reflexive, symmetrical and transitive. The equivalence relation  $R$  partitions the set  $U$  into disjoint subsets. Elements in the same equivalence class are said to be indistinguishable. Equivalence classes of  $R$  are called elementary sets. Every union of elementary sets is called a definable set [1]. The empty set is considered to be a definable set, thus all the definable sets form a Boolean algebra.  $(U, R)$  is called an approximation space. Given an arbitrary set  $X \subseteq U$ , one can characterize  $X$  by a pair of lower and upper approximations. The lower approximation  $\underline{apr}_R X$  is the greatest definable set contained in  $X$ , and the upper approximation  $\overline{apr}_R X$  is the least definable set containing  $X$ . They can be computed by two equivalent formulae

$$\begin{aligned}\underline{apr}_R X &= \{x : [x]_R \subseteq X\}, & \overline{apr}_R X &= \{x : [x]_R \cap X \neq \emptyset\}, \\ \underline{apr}_R X &= \bigcup \{[x]_R : [x]_R \subseteq X\}, & \overline{apr}_R X &= \bigcup \{[x]_R : [x]_R \cap X \neq \emptyset\}.\end{aligned}$$

Let  $\mathbf{R}$  be a family of equivalence relations and let  $A \in \mathbf{R}$ , denote  $\text{IND}(\mathbf{R}) = \bigcap \{R : R \in \mathbf{R}\}$ . We say that  $A$  is dispensable in  $\mathbf{R}$  if  $\text{IND}(\mathbf{R}) = \text{IND}(\mathbf{R} - \{A\})$ ; otherwise  $A$  is indispensable in  $\mathbf{R}$ . The family  $\mathbf{R}$  is independent

if each  $A \in \mathbf{R}$  is indispensable in  $\mathbf{R}$ ; otherwise  $\mathbf{R}$  is dependent.  $\mathbf{Q} \subseteq \mathbf{P}$  is a reduct of  $\mathbf{P}$  if  $\mathbf{Q}$  is independent and  $\text{IND}(\mathbf{Q}) = \text{IND}(\mathbf{P})$ . The set of all indispensable relations in  $\mathbf{P}$  will be called the core of  $\mathbf{P}$ , and will be denoted as  $\text{CORE}(\mathbf{P})$ . Clearly,  $\text{CORE}(\mathbf{P}) = \cap \text{RED}(\mathbf{P})$ , where  $\text{RED}(\mathbf{P})$  is the family of all reducts of  $\mathbf{P}$ . In [13] the discernibility matrix method was presented to compute all the reducts of information systems and relative reducts of decision systems.

### 3. Covers and covering generalized rough sets

In this section we first give some approaches and examples to construct covers from data set. These approaches and examples can be viewed as the applied background of covering rough sets. We then conclude the lower and upper approximations of covering generalized rough sets.

**Definition 3.1** ([10]). Let  $U$  be a finite universe of discourse,  $\mathbf{C}$  a family of subsets of  $U$ . If none of the subsets in  $\mathbf{C}$  is empty and  $\cup \mathbf{C} = U$ ,  $\mathbf{C}$  is called a cover of  $U$ .

It is clear that a partition of  $U$  is certainly a cover of  $U$ , and the concept of a cover is an extension of the concept of a partition.

The existing researches on covering generalized rough sets are mainly concentrated on the approximation operators. Less effort has been put on the application of covering generalized rough sets. First of all, suitable data sets and methods of constructing covers from data sets should be adopted for the application of covering generalized rough sets before we study the covering generalized rough sets. We first present some suitable data sets for covering generalized rough sets.

The information system with multiattribute values for an attribute for some objects is one suitable data set for covering generalized rough sets. This kind of data set is available when we are not sure the attribute values of an attribute for some objects. So we have to list all the possible attribute values. One example of this kind of data set is the combination of several information systems. This idea will be illustrated by the following example:

**Example 3.2.** Let us consider an evaluation problem of credit card applicants. Suppose  $U = \{x_1, \dots, x_9\}$  is a set of nine applicants,  $E = \{\text{education; salary}\}$  is a set of two attributes, the values of “education” are {best; better; good}, and the values of “salary” are {high; middle; low}. We have three specialists  $\{A, B, C\}$  to evaluate the attributes of these applicants. It is possible that their evaluation results to the same attributes values are not the same to one another. The evaluation results are listed below.

For attribute “education”:

$A$  : best =  $\{x_1, x_4, x_5, x_7\}$ , better =  $\{x_2, x_8\}$ , good =  $\{x_3, x_6, x_9\}$ ;

$B$  : best =  $\{x_1, x_2, x_4, x_7, x_8\}$ , better =  $\{x_5\}$ , good =  $\{x_3, x_6, x_9\}$ ;

$C$  : best =  $\{x_1, x_4, x_7\}$ , better =  $\{x_2, x_8\}$ , good =  $\{x_3, x_5, x_6, x_9\}$ .

For attribute “salary”:

$A$  : high =  $\{x_1, x_2, x_3\}$ , middle =  $\{x_4, x_5, x_6, x_7, x_8\}$ , low =  $\{x_9\}$ ;

$B$  : high =  $\{x_1, x_2, x_3\}$ , middle =  $\{x_4, x_5, x_6, x_7\}$ , low =  $\{x_8, x_9\}$ ;

$C$  : high =  $\{x_1, x_2, x_3\}$ , middle =  $\{x_4, x_5, x_6, x_8\}$ , low =  $\{x_7, x_9\}$ .

Assume the evaluations of these specialists are of the same importance. If we want to combine these evaluations together without losing information, we should unite the evaluations given by each specialist for every attribute value. We have the classification as in Table 1.

This classification is a cover and not a partition, and reflects a kind of uncertainty caused by differences in interpretation of the data. The covering rough set can be applied to extract valuable rules from this kind of information system.

Another kind of dataset suitable for covering generalized rough sets is fuzzy data set [20]. Suppose  $U$  is a universe,  $C$  is a set of  $N$  fuzzy attributes. Every sample  $x_i$  in  $U$  is described by the fuzzy attributes in  $C$ , that is, they have been measured as partial membership degrees,  $\mu_i^j$  ( $j = 1, \dots, N$ ), which are graded in the interval  $[0, 1]$ :

Table 1

Salary	Education		
	Best	better	good
High	$\{x_1, x_2\}$	$\{x_2\}$	$\{x_3\}$
Middle	$\{x_4, x_5, x_7, x_8\}$	$\{x_5, x_8\}$	$\{x_5, x_6\}$
Low	$\{x_7, x_8\}$	$\{x_8\}$	$\{x_9\}$

Table 2

	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$	$C_6$
$x_1$	1	0	0	1	0	0
$x_2$	1/3	2/3	0	1	0	0
$x_3$	0	0	1	1	0	0
$x_4$	1	0	0	0	1	0
$x_5$	1/3	1/3	1/3	0	1	0
$x_6$	0	0	1	0	1	0
$x_7$	1	0	0	0	2/3	1/3
$x_8$	1/3	2/3	0	0	2/3	1/3
$x_9$	0	0	1	0	0	1

$x_i = [\mu_i^1, \mu_i^2, \dots, \mu_i^N]$ . A similarity relation  $R$  on the universe  $U$  can be defined using these fuzzy attributes [20]. This similarity relation  $R$  is a fuzzy relation satisfying reflexivity, symmetry and  $T$ -transitivity, here  $T$  is a triangular norm. The fuzzy set  $[x]_R$ , defined as  $[x]_R(y) = R(x, y), \forall y \in U$ , is called the fuzzy similarity class of  $x \in U$ . By taking  $\alpha \in (0, 1]$ , then the collection of all the  $\alpha$ -cut of  $[x]_R$ , i.e.,  $\{([x]_R)_\alpha : x \in U\}$ , is a cover of  $U$ . Here  $([x]_R)_\alpha = \{y \in U : [x]_R(y) = R(x, y) \geq \alpha\}$ .  $\alpha$  is a criterion of similarity that can be set beforehand.  $R(x, y) \geq \alpha$  means  $y$  is similar to  $x$ , otherwise  $y$  is not similar to  $x$ . We have the following example to illustrate our idea:

**Example 3.3.** In Example 3.2 if every applicant is described by six fuzzy attributes:  $C_1$  = best education,  $C_2$  = better education,  $C_3$  = good education,  $C_4$  = high salary,  $C_5$  = middle salary and  $C_6$  = low salary, then we can compute the membership degrees of every applicant by the evaluations of three specialists  $\{A, B, C\}$ , for instance, two specialists believe  $x_2$  is with better education, then  $C_2(x_2) = 2/3$ . We have Table 2.

The fuzzy similarity relation  $R$  can be computed by  $R(x_i, x_j) = \inf_{1 \leq k \leq 6} (1 - |\mu_i^k - \mu_j^k|)$ , as follows:

$$(R(x_i, x_j)) = \begin{pmatrix} 1 & 1/3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 1 & 1/3 & 0 & 2/3 & 1/3 & 0 \\ & & & & 1 & 1/3 & 1/3 & 2/3 & 0 \\ & & & & & 1 & 0 & 0 & 0 \\ & & & & & & 1 & 1/3 & 0 \\ & & & & & & & 1 & 0 \\ & & & & & & & & 1 \end{pmatrix}.$$

$R$  is a fuzzy  $T_L$ -similarity relation,  $T_L$  is the Lukasiewicz  $t$ -norm defined as  $T_L(x, y) = \max\{0, x + y - 1\}$ . By taking  $\alpha = 1/3$ , then we can get a cover as

$$\{\{x_1, x_2\}, \{x_3\}, \{x_4, x_5, x_7, x_8\}, \{x_4, x_5, x_6, x_7, x_8\}, \{x_5, x_6\}, \{x_9\}\}.$$

Covering generalized rough sets are also suitable for rough sets with respect to similarity relation [9]. Suppose  $R$  is a reflexivity relation on  $U$ , denote  $R_s(x) = \{y : (x, y) \in R\}$ , then  $\{R_s(x) : x \in U\}$  is also a cover of  $U$ . So the rough sets with respect to similarity relation in [9] can be brought into the framework of covering rough sets theory from the theoretical viewpoint. Generally speaking, a cover cannot be obtained by this method, i.e. if  $C$  is a cover, it will not always be true that there exists a reflexivity relation  $R$  such that  $C = \{R_s(x) : x \in U\}$ .

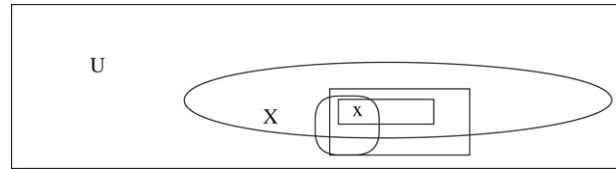


Fig. 1.

Now we present an applied background for covering generalized rough sets. Certainly the covering generalized rough sets are not just suitable for these applications. We should study the approximations in covering generalized rough sets in order to extract rules from data sets.

The existing research on covering generalized rough sets mainly focuses on the approximation operators. While the ideas of constructing lower approximation operator are similar, the ideas of defining upper approximation operator are different. These important observations have been studied in detail in [21] under a wider framework of complete completely distributive lattice. We only list the definitions of lower and upper approximations in covering generalized rough sets and discuss the upper approximations further as follows:

**Definition 3.4** ([10]). The ordered pair  $(U, \mathbf{C})$  is called a covering approximation space, where  $U$  is any nonempty set called a universe, and  $\mathbf{C}$  is its finite cover.

**Definition 3.5** ([10]). Let  $(U, \mathbf{C})$  be the approximation space,  $x \in U$ . The following family  $Md(x) = \{K \in \mathbf{C} : x \in K \wedge \forall S \in \mathbf{C}(x \in S \wedge S \subseteq K \Rightarrow K = S)\}$  is called the minimal description of the object  $x$ . In another words, every element in  $Md(x)$  is a minimal one including  $x$ .

**Definition 3.4** ([10]). For any  $X \subseteq U$ , the set  $X_* = \cup\{K \in \mathbf{C} : K \subseteq X\}$  is called the lower approximation of the set  $X$ .

**Definition 3.5.** For any  $X \subseteq U$ , the set  $X^* = \cup\{K \in Md(x) : x \in X\}$  is called the upper approximation of the set  $X$ .

**Definition 3.5** is the special case of the upper approximation defined on a complete completely distributive lattice in [21]. **Definition 3.5** is different from the definitions of upper approximation in [10,11]. In [10] the upper approximation of  $X$  is defined as  $X_* \cup \{K \in Md(x) : x \in X - X_*\}$  and in [11] the upper approximation of  $X$  is defined as  $\{y \in U : \exists C \in \mathbf{C}, y \in C \text{ and } C \cap X \neq \emptyset\}$ . In [21] we have argued that it is possible to lose some useful information by the definition of upper approximation in [10] and it is also possible to include some unnecessary information by the definition of upper approximation in [11], a reasonable definition of upper approximation should maintain the minimal and most informative description of  $X$ . Our **Definition 3.5** is a reasonable one.

**Fig. 1** will illustrate our idea.

In **Fig. 1**, suppose  $U$  is a finite universe and let  $\mathbf{C}$  be a cover of  $U$ ,  $X$  (the ellipse) is a subset of  $U$ , the big, small and round rectangles are all elements in  $\mathbf{C}$  which include  $x \in X$ . Only the small rectangle, all the three rectangles, and both of the small rectangle and the round rectangle are considered respectively when computing the upper approximation of  $X$ , by the definition of upper approximation in [10], the definition of upper approximation in [11] and **Definition 3.5**.

However, one important thing we should point out is that **Definition 3.5** is just suitable for the case of finite universe. It is possible that  $Md(x) = \emptyset$  for every  $x$  in the infinite universe. The following example will illustrate this statement.

**Example 3.6.** Suppose  $U = (0, 1)$ . For every  $x \in U$ , take a sequence  $\{\varepsilon_n^x\} \subseteq U$  such that  $\varepsilon_n^x \rightarrow 0$ ,  $\{\varepsilon_n^x\}$  is strictly decreasing and  $(x - \varepsilon_1^x, x + \varepsilon_1^x) \subseteq U$ . Then  $\{(x - \varepsilon_n^x, x + \varepsilon_n^x) : x \in U, n = 1, 2, \dots\}$  is a cover of  $U = (0, 1)$  and for every  $x \in U$ ,  $Md(x) = \emptyset$ .

To extract rules from data set with multiple attributes by **Definitions 3.4** and **3.5**, we should study intersection of several covers. We discuss this topic in the next section.

#### 4. On the reduct of covering generalized rough sets

It is well known that attributes reduct is the most important result in rough set theory to distinguish itself from other theories. For any generalization of rough set theory, the research on attributes reduct is available and necessary.

In [12] it was claimed that the reduct of covering generalized rough sets had been studied, but their reduct is to reduce “redundant” members in a fixed cover, which does not agree with the original meaning of reduct using traditional rough sets. For a family of covers, how to determine what should be invariant after reduct, i.e. how to decide which cover is dispensable is the key problem for the reduct of covering generalized rough sets. In this section first we present a method to determine dispensable covers by defining induced cover and their interaction such that the reduct of a family of covers is naturally defined, then we study relations between two different elements with respect to a family of covers and by using these relations we propose the discernibility matrix method to compute all the reducts of a family of covers. Our method is the generalization of the method in [13] for computing reducts of traditional rough sets theory.

For two elements in a cover, they may have nonempty overlap. This means that if a cover is employed to express an attribute and the elements in this cover express the attributes values, then the attribute values may have nonempty overlap. For the objects in the nonempty overlapping part, one possible description is to employ both of these two attributes values so that no information is lost. For example, for the attribute “age” with the attribute values “young”, “middle” and “old”, a 35 year-old person may be young man or middle age, the most complete description to this person is ‘young or middle age’ for losing no information. By this way, we can get the most complete description of every object in the universe by defining the induced cover of a given cover.

**Definition 4.1.** Suppose  $U$  is a finite universe and  $\mathbf{C} = \{C_1, C_2, \dots, C_n\}$  is a cover of  $U$ . For every  $x \in U$ , let  $C_x = \cap\{C_j : C_j \in \mathbf{C}, x \in C_j\}$ , then  $\text{Cov}(\mathbf{C}) = \{C_x : x \in U\}$  is also a cover of  $U$ , we call it the induced cover of  $\mathbf{C}$ .

For every  $x \in U$ ,  $C_x$  is the minimal set in  $\text{Cov}(\mathbf{C})$  including  $x$ , i.e.  $C_x$  is the most complete description of  $x$  with respect to  $\mathbf{C}$ . If  $\mathbf{C}$  is an attribute and every  $C_i$  denotes an attribute value of  $\mathbf{C}$ , i.e. the collection of objects in  $U$  takes certain attribute value. Suppose  $C_x = C_1 \cap C_2$ , this implies the possible values of  $x$  with respect to  $\mathbf{C}$  are  $C_1$  or  $C_2$ , i.e. the relation between  $C_1$  and  $C_2$  is disjunctive.

However, Definition 4.1 is not suitable for the incomplete information system found in [3,19]. In an incomplete information system, if an object losses its attribute value, then it is believed every value of this attribute is available for this object [3,19]. All the objects with missing attribute value will be classified as a single class by Definition 4.1. This statement can be easily examined with the example in [3] and we omit this discussion.

$\text{Cov}(\mathbf{C}) = \mathbf{C}$  if and only if  $\mathbf{C}$  is a partition. For every  $x, y \in U$ , if  $y \in C_x$  then  $C_x \supseteq C_y$ , so if  $y \in C_x$  and  $x \in C_y$ , then  $C_x = C_y$ . Every element in  $\text{Cov}(\mathbf{C})$  cannot be written as the union of other elements in  $\text{Cov}(\mathbf{C})$ . For  $\text{Cov}(\mathbf{C})$  in every  $Md(x)$  there is only one element, and the upper approximation obtained by  $\text{Cov}(\mathbf{C})$  is smaller than the one obtained by  $\mathbf{C}$  with Definition 3.5, while the lower approximation is bigger, this implies we can extract more precise rules with  $\text{Cov}(\mathbf{C})$  than with  $\mathbf{C}$ .

**Definition 4.2.** Suppose  $U$  is a finite universe and  $\Delta = \{\mathbf{C}_i : i = 1, \dots, m\}$  is a family of covers of  $U$ . For every  $x \in U$ , let  $\Delta_x = \cap\{C_{ix} : C_{ix} \in \text{Cov}(\mathbf{C}_i), x \in C_{ix}\}$ , then  $\text{Cov}(\Delta) = \{\Delta_x : x \in U\}$  is also a cover of  $U$ . We call it the induced cover of  $\Delta$ .

Clearly  $\Delta_x$  is the intersection of all the elements in every  $\mathbf{C}_i$  which includes  $x$ . This implies that  $\Delta_x$  is the minimal set in  $\text{Cov}(\Delta)$  including  $x$  for every  $x \in U$ , i.e., the most complete description of  $x$  with respect to  $\Delta$ .  $\text{Cov}(\Delta)$  can be viewed as the intersection of covers in  $\Delta$ , it is the final classification of the universe  $U$  by covers in  $\Delta$  and may not be a partition.  $\Delta_x = \cap\{C_{ix} : C_{ix} \in \text{Cov}(\mathbf{C}_i), x \in C_{ix}\}$  means the relation among  $C_{ix}$  is conjunctive. For example, in Table 1 of Example 3.2,  $\{x_2\} = \{x_1, x_2\} \cap \{x_2\}$  implies the description of  $x_2$  is “(best education  $\vee$  better education)  $\wedge$  (high salary)”, i.e. the relation between “best education” and “better education” is disjunctive and the relation between “best education  $\vee$  better education” and “high salary” is conjunctive. If  $x_2$  gets his credit card permission, then a rule as “(best education  $\vee$  better education)  $\wedge$  (high salary)  $\Rightarrow$  credit card permission” can be extracted.

If every cover in  $\Delta$  is a partition, then  $\text{Cov}(\Delta)$  is also a partition and  $\Delta_x$  is the equivalence class including  $x$ . For every  $x, y \in U$ , if  $y \in \Delta_x$ , then  $\Delta_x \supseteq \Delta_y$ , so if  $y \in \Delta_x$  and  $x \in \Delta_y$ , then  $\Delta_x = \Delta_y$ . Every element in  $\text{Cov}(\Delta)$  cannot be written as the union of other elements in  $\text{Cov}(\Delta)$ . For every  $\mathbf{C}_i \in \Delta$ , if  $\text{Cov}(\Delta - \{\mathbf{C}_i\}) = \text{Cov}(\Delta)$ , then  $\mathbf{C}_i$  is called dispensable in  $\Delta$ , otherwise  $\mathbf{C}_i$  is called indispensable in  $\Delta$ . Here  $\text{Cov}(\Delta - \{\mathbf{C}_i\}) = \text{Cov}(\Delta)$  means for every  $x \in U$ ,  $\Delta_x = \Delta'_x$  holds where  $\Delta' = \Delta - \{\mathbf{C}_i\}$ . For every  $\mathbf{P} \subseteq \Delta$ , if every element in  $\mathbf{P}$  is indispensable and  $\text{Cov}(\mathbf{P}) = \text{Cov}(\Delta)$ , then  $\mathbf{P}$  is called the reduct of  $\Delta$ . The collection of all the indispensable elements in  $\Delta$  is called the core of  $\Delta$ , denoted as  $\text{Core}(\Delta)$ . Similar to the traditional rough set theory it can be proved that the core of  $\Delta$  is the intersection of all the reducts of  $\Delta$ .



In traditional rough set theory for every  $x, y \in U$ , the two equivalence classes including these two objects are either equal to each other or have an empty overlap. If their equivalence classes are equal, then these two objects are called indiscernible. Based on this statement, the method of discernibility matrix to compute all the reducts of traditional rough sets was presented in [13]. However, for the covering generalized rough sets, things are quite different and more complex. For every  $x, y \in U$ , there are three possible relations with respect to  $\Delta$  between  $x$  and  $y$ :  $R(1) \Delta_x = \Delta_y$ ;  $R(2) \Delta_x \subset \Delta_y$  or  $\Delta_x \supset \Delta_y$ ;  $R(3) \Delta_x \not\subset \Delta_y, \Delta_y \not\subset \Delta_x$  i.e.,  $\Delta_x$  and  $\Delta_y$  cannot be included by each other. If  $\text{Cov}(\Delta)$  is a partition, then these three relations are just the case in traditional rough sets as mentioned above. In the following we call these relations the original relations between  $x$  and  $y$  with respect to  $\Delta$ . We have the following statements to characterize these three relations respectively:

**Proposition 4.3.** Suppose  $U$  is a finite universe and  $\Delta = \{C_i : i = 1, \dots, m\}$  is a family of covers of  $U$ .

- (1)  $\Delta_x = \Delta_y$  if and only if for every  $C_i \in \Delta$  we have  $C_{ix} = C_{iy}$ .
- (2)  $\Delta_x \supset \Delta_y$  if and only if for every  $C_i \in \Delta$  we have  $C_{ix} \supseteq C_{iy}$  and there is a  $C_{i_0} \in \Delta$  such that  $C_{i_0x} \supset C_{i_0y}$ .
- (3)  $\Delta_x \not\subset \Delta_y$  and  $\Delta_y \not\subset \Delta_x$  hold if and only if there are  $C_i, C_j \in \Delta$  such that  $C_{ix} \subset C_{iy}$  and  $C_{jx} \supset C_{jy}$  or there is a  $C_{i_0} \in \Delta$  such that  $C_{i_0x} \not\subset C_{i_0y}$  and  $C_{i_0y} \not\subset C_{i_0x}$ .

**Proof.** (1) If for every  $C_i \in \Delta$  we have  $C_{ix} = C_{iy}$ , clearly  $\Delta_x = \Delta_y$  holds.

If there exists a  $C_{i_0} \in \Delta$  such that  $C_{i_0x} \neq C_{i_0y}$ , then at least one of  $x \notin C_{i_0y}$  and holds, this implies  $\Delta_x \neq \Delta_y$ .

(2) If for every  $C_i \in \Delta$  we have  $C_{ix} \supseteq C_{iy}$ , then  $\Delta_x \supseteq \Delta_y$  holds. If there is a  $C_{i_0} \in \Delta$  such that  $C_{i_0x} \supset C_{i_0y}$  holds, then we have  $x \notin C_{i_0y}$ , this statement implies  $\Delta_x \supset \Delta_y$ .

If  $\Delta_x \supset \Delta_y$ , then we have  $y \in \Delta_x$ . So for every  $C_i \in \Delta$  we have  $y \in C_{ix}$  which implies  $C_{ix} \supseteq C_{iy}$ . By  $\Delta_x \supset \Delta_y$  we have  $x \notin \Delta_y$ , so there is a  $C_{i_0} \in \Delta$  such that  $x \notin C_{i_0y}$ , so  $C_{i_0x} \supset C_{i_0y}$ .

(3) If there are  $C_i, C_j \in \Delta$  such that  $C_{ix} \subset C_{iy}$  and  $C_{jx} \supset C_{jy}$ , then  $y \notin C_{ix}$  which implies  $y \notin \Delta_x$  and  $x \notin C_{jy}$  which implies  $x \notin \Delta_y$ , thus we have  $\Delta_x \not\subset \Delta_y$  and  $\Delta_y \not\subset \Delta_x$ .

If there is a  $C_{i_0} \in \Delta$  such that  $C_{i_0x} \not\subset C_{i_0y}$  and  $C_{i_0y} \not\subset C_{i_0x}$ , then we have  $x \notin C_{i_0y}$  and  $y \notin C_{i_0x}$ , so  $x \notin \Delta_y$  and  $y \notin \Delta_x$  which implies  $\Delta_x \not\subset \Delta_y$  and  $\Delta_y \not\subset \Delta_x$ .

If  $\Delta_x \not\subset \Delta_y$  and  $\Delta_y \not\subset \Delta_x$  hold and there is no  $C_{i_0} \in \Delta$  such that  $C_{i_0x} \not\subset C_{i_0y}$  and  $C_{i_0y} \not\subset C_{i_0x}$ , then by the proof of (2) there must have  $C_i, C_j \in \Delta$  such that  $C_{ix} \subset C_{iy}$  and  $C_{jx} \supset C_{jy}$ , otherwise it will lead to  $\Delta_x \supset \Delta_y$  or  $\Delta_x \subset \Delta_y$  which is a contradiction.

By the above three relations we can have the following theorem to determine which cover is indispensable in  $\Delta$ :

□

**Theorem 4.4.** Suppose  $U$  is a finite universe and  $\Delta = \{C_i : i = 1, \dots, m\}$  is a family of covers of  $U$ ,  $C_i \in \Delta$ . Then  $\text{Cov}(\Delta - \{C_i\}) \neq \text{Cov}(\Delta)$  if and only if there is at least a pair of  $x, y \in U$  whose original relation with respect to  $\Delta$  is changed after deleting  $C_i$  from  $\Delta$ .

**Proof.** We denote  $\text{Cov}(\Delta - \{C_i\}) = \{\Delta'_x : x \in U\}$ . If  $\text{Cov}(\Delta - \{C_i\}) \neq \text{Cov}(\Delta)$ , then there are  $x_0, y_0 \in U$  such that  $y_0 \in \Delta'_{x_0}$  and  $y_0 \notin \Delta_{x_0}$ , which imply  $\Delta'_{y_0} \subseteq \Delta'_{x_0}$  and  $\Delta_{y_0} \not\subset \Delta_{x_0}$ . So the original relation of  $x_0, y_0$  with respect to  $\Delta$  is changed after deleting  $C_i$  from  $\Delta$ .

For any two elements  $x, y \in U$ ,  $\Delta_x = \Delta_y$  if and only if for every  $C_i \in \Delta$  we have  $C_{ix} = C_{iy}$ . So if we delete an element from  $\Delta$  and the original relation of  $x, y \in U$  with respect to  $\Delta$  is changed, then  $\Delta_x = \Delta_y$  could not hold.

Suppose we delete  $C_i$  from  $\Delta$  and the original relation of  $x, y \in U$  with respect to  $\Delta$  is changed. Then at least one of  $\Delta_x \neq \Delta'_x$  and  $\Delta_y \neq \Delta'_y$  holds which implies  $\text{Cov}(\Delta - \{C_i\}) \neq \text{Cov}(\Delta)$ .

The purpose of reduct of  $\Delta$  is to find the minimal subset of  $\Delta$  to keep every element in  $\text{Cov}(\Delta)$  invariant. By Theorem 4.4 we know that it is equivalent to finding the minimal subset of  $\Delta$  by keeping relations between every two elements in the universe invariant. In [13] the algorithm employing the discernibility matrix was proposed to compute the reducts of rough sets. In the following we generalize it to compute the reducts of covering generalized rough sets. Suppose  $U = \{x_1, \dots, x_n\}$ . By  $M(U, \Delta)$  we denote an  $n \times n$  matrix  $(c_{ij})$ , called the discernibility matrix of  $(U, \Delta)$ , such that

$$c_{ij} = \begin{cases} \phi, & \Delta_{x_i} = \Delta_{x_j}; \\ \{C \in \Delta : C_{x_i} \subset C_{x_j}\}, & \Delta_{x_i} \subset \Delta_{x_j} \text{ or } \{C \in \Delta : C_{x_{ji}} \subset C_{x_i}\}, & \Delta_{x_j} \subset \Delta_{x_i}; \\ \{C \in \Delta : (C_{x_i} \not\subset C_{x_j}) \wedge (C_{x_j} \not\subset C_{x_i})\} \cup \{C_s \wedge C_t : (C_{sx_i} \subset C_{sx_j}) \wedge (C_{tx_j} \subset C_{tx_i})\}, & \Delta_{x_i} \not\subset \Delta_{x_j} \wedge \Delta_{x_j} \not\subset \Delta_{x_i} \end{cases}$$

for  $x_i, x_j \in U$ .

Since  $M(U, \Delta)$  is symmetrical and  $c_{ii} = \phi$  for  $i = 1, \dots, n$ , we represent  $M(U, \Delta)$  only by elements in the lower triangle of  $M(U, \Delta)$ , i.e. the  $c_{ij}$ 's with  $1 \leq j < i \leq n$ .

A discernibility function  $f(U, \Delta)$  for  $(U, \Delta)$  is a Boolean function of  $m$  Boolean variables  $\overline{C_1}, \dots, \overline{C_m}$  corresponding to the covers  $C_1, \dots, C_m$ , respectively, and is defined as follows:

$$f(U, \Delta)(\overline{C_1}, \dots, \overline{C_m}) = \wedge \{ \vee(c_{ij}) : 1 \leq j < i \leq n, c_{ij} \neq \phi \}$$

where  $\vee(c_{ij})$  is the disjunction of all elements in  $c_{ij}$  as  $C$  or  $C_s \wedge C_t$ . We have the following proposition for the core.  $\square$

**Proposition 4.5.**  $\text{Core}(\Delta) = \{C \in \Delta : c_{ij} = \{C\} \vee c_{ij} = \{C \wedge C_t : t = 1, \dots, k\} \text{ for some } i, j\}$ .

**Proof.** Suppose  $C \in \text{Core}(\Delta)$ , then  $\text{Cov}(\Delta - \{C\}) \neq \text{Cov}(\Delta)$ , this implies there exist  $x_i$  and  $x_j$  whose original relation with respect to  $\Delta$  are changed after deleting  $C$  from  $\Delta$ . If  $x_i$  and  $x_j$  are with relation  $R(2)$  with respect to  $\Delta$ , then  $c_{ij} = \{C\}$ . If  $x_i$  and  $x_j$  are with relation  $R(3)$  with respect to  $\Delta$ , then  $c_{ij} = \{C\}$  or there exists  $\{C_t : t = 1, \dots, k\} \subseteq \Delta$  such that  $c_{ij} = \{C \wedge C_t : t = 1, \dots, k\}$ .

If  $c_{ij} = \{C\}$  or  $c_{ij} = \{C \wedge C_t : t = 1, \dots, k\}$  for some  $i, j$ , it is clear that  $C \in \text{Core}(\Delta)$ . Hence we complete the proof.  $\square$

**Proposition 4.6.** Suppose  $\Delta' \subset \Delta$ , then  $\text{Cov}(\Delta') = \text{Cov}(\Delta)$  if and only if  $\Delta' \cap c_{ij} \neq \phi$  for every  $c_{ij} \neq \phi$ . Here  $C_s \wedge C_t \in c_{ij}$  belonging to  $\Delta' \cap c_{ij}$  means  $\{C_s, C_t\} \subseteq \Delta'$ .

**Proof.** If  $\text{Cov}(\Delta') = \text{Cov}(\Delta)$ , then  $\Delta'$  contains a reduct of  $(U, \Delta)$ , then any two objects with relation  $R(2)$  or  $R(3)$  with respect to  $\Delta$  are also with relation  $R(2)$  or  $R(3)$  with respect to  $\Delta'$ . Hence if  $c_{ij} \neq \phi$ , then  $\Delta' \cap c_{ij} \neq \phi$ .

If  $\Delta' \cap c_{ij} \neq \phi$  for  $c_{ij} \neq \phi$ , then it means that in  $\Delta'$  we have enough covers to keep the original relations with respect to  $\Delta$  between all the objects in  $U$ , i.e.  $\Delta'$  contains a reduct of  $(U, \Delta)$ , which implies  $\text{Cov}(\Delta') = \text{Cov}(\Delta)$ .  $\square$

**Corollary 4.7.** Suppose  $\Delta' \subset \Delta$ , then  $\Delta'$  is a reduct of  $\Delta$  if and only if it is the minimal set satisfying  $\Delta' \cap c_{ij} \neq \phi$  for  $c_{ij} \neq \phi$ .

Let  $g(U, \Delta)$  be the reduced disjunctive form of  $f(U, \Delta)$  obtained from  $f(U, \Delta)$  by applying the multiplication and absorption laws as many times as possible. Then there exist  $l$  and  $\Delta_k \subseteq \Delta$  for  $k = 1, \dots, l$  such that  $g(U, \Delta) = (\wedge \Delta_1) \vee \dots \vee (\wedge \Delta_l)$  where every element in  $\Delta_k$  only appears one time. Let  $\text{RED}(\Delta)$  be the collection of all the reducts of  $(U, \Delta)$ . We have the following theorem:

**Theorem 4.8.**  $\text{RED}(\Delta) = \{\Delta_1, \dots, \Delta_l\}$ .

**Proof.** For every  $k = 1, \dots, l$ , we have  $\wedge \Delta_k \leq \vee c_{ij}$ , so  $\Delta_k \cap c_{ij} \neq \phi$  if  $c_{ij} \neq \phi$ . Let  $\Delta'_k = \Delta_k - \{C\}$ , then  $g(U, \Delta) \neq \vee_{r=1}^{k-1} (\wedge \Delta_r) \vee (\wedge \Delta'_k) \vee (\vee_{r=k+1}^l \Delta_r)$  and  $g(U, \Delta) < \vee_{r=1}^{k-1} (\wedge \Delta_r) \vee (\wedge \Delta'_k) \vee (\vee_{r=k+1}^l \Delta_r)$ . If for every  $c_{ij} \neq \phi$  we have  $\Delta'_k \cap c_{ij} \neq \phi$ , then  $\wedge \Delta'_k \leq \vee c_{ij}$  for every  $c_{ij} \neq \phi$ . This implies  $g(U, \Delta) \geq \vee_{r=1}^{k-1} (\wedge \Delta_r) \vee (\wedge \Delta'_k) \vee (\vee_{r=k+1}^l \Delta_r)$  and  $g(U, \Delta) = \vee_{r=1}^{k-1} (\wedge \Delta_r) \vee (\wedge \Delta'_k) \vee (\vee_{r=k+1}^l \Delta_r)$  which is a contradiction. Hence there exists  $c_{i_0 j_0} \neq \phi$  such that  $\Delta'_k \cap c_{i_0 j_0} = \phi$  which implies  $\Delta_k$  is a reduct of  $(U, \Delta)$ .

For every  $X \in \text{RED}(\Delta)$ , we have  $X \cap c_{ij} \neq \phi$  for every  $c_{ij} \neq \phi$ , so we have  $f(U, \Delta) \wedge (\wedge X) = \wedge (\vee c_{ij}) \wedge (\wedge X) = \wedge X$ , this implies  $\wedge X \leq f(U, \Delta) = g(U, \Delta)$ . Suppose for every  $k$  we have  $\Delta_k - X \neq \phi$ , then for every  $k$  one can find  $C_k \in \Delta_k - X$ . By rewriting  $g(U, \Delta) = (\vee_{k=1}^l C_k) \wedge \Phi$ , then we have  $\wedge X \leq \vee_{k=1}^l C_k$ . So there is  $C_{k_0}$  such that  $\wedge X \leq C_{k_0}$ , this implies  $C_{k_0} \in X$  which is a contradiction. So  $\Delta_{k_0} \subseteq X$  for some  $k_0$ , since both  $X$  and  $\Delta_{k_0}$  are reducts, we have  $X = \Delta_{k_0}$ . Hence  $\text{RED}(\Delta) = \{\Delta_1, \dots, \Delta_l\}$ .

It should be mentioned that if every cover is a partition, then our method for computing the reducts is coincided with the method of computing reducts of rough sets in [13]. In [13] it was pointed out that the computing of the reducts of traditional rough sets is NP hard, and it is just the special case of computing reducts of covering generalized rough sets in this paper. So the computing of reducts of covering generalized rough sets is at least NP hard. From the theoretical viewpoint of our method we can compute all the reducts of the covering generalized rough sets. Our research on this topic plays as an important theoretical role as the research in [13] in the traditional rough set theory. For some covering generalized rough sets with additional conditions it is possible to design efficient algorithms based



on the method in this section, which will be our future work. Our idea in this section could be illustrated by the following simple application:  $\square$

**Example 4.9.** Now we consider a house evaluation problem. Suppose  $U = \{x_1, \dots, x_9\}$  be a set of nine houses,  $E = \{\text{price; color; structure; surrounding}\}$  be a set of attributes, the values of “price” are {high; middle; low}, the values of “colour” are {good; bad}, the values of “structure” are {reasonable; ordinary; unreasonable}, and the values of “surrounding” are {quiet; a little noisy; noisy; quite noisy}. We have four specialists to evaluate the attributes of these houses, they are  $\{A, B, C, D\}$ , then it is possible that their evaluation results in the same attribute values are not the same as one another. The evaluation results are listed below.

For attribute “price”

$A$  : high =  $\{x_1, x_4, x_5, x_7\}$ , middle =  $\{x_2, x_8\}$ , low =  $\{x_3, x_6, x_9\}$ ;  
 $B$  : high =  $\{x_1, x_2, x_4, x_7, x_8\}$ , middle =  $\{x_5\}$ , low =  $\{x_3, x_6, x_9\}$ ;  
 $C$  : high =  $\{x_1, x_4, x_7\}$ , middle =  $\{x_8\}$ , low =  $\{x_2, x_3, x_5, x_6, x_9\}$ ;  
 $D$  : high =  $\{x_1, x_4, x_7\}$ , middle =  $\{x_5\}$ , low =  $\{x_2, x_3, x_6, x_8, x_9\}$ .

For attribute “color”:

$A$  : good =  $\{x_1, x_2, x_3, x_6\}$ , bad =  $\{x_4, x_5, x_7, x_8, x_9\}$ ;  
 $B$  : good =  $\{x_1, x_2, x_3, x_5\}$ , bad =  $\{x_4, x_6, x_7, x_8, x_9\}$ ;  
 $C$  : good =  $\{x_1, x_2, x_3, x_4\}$ , bad =  $\{x_5, x_6, x_7, x_8, x_9\}$ ;  
 $D$  : good =  $\{x_1, x_2, x_3\}$ , bad =  $\{x_4, x_5, x_6, x_7, x_8, x_9\}$ .

For attribute “structure”:

$A$  : reasonable =  $\{x_1, x_2, x_3\}$ , ordinary =  $\{x_4, x_5, x_6, x_7, x_8\}$ , unreasonable =  $\{x_9\}$ ;  
 $B$  : reasonable =  $\{x_1, x_2, x_3\}$ , ordinary =  $\{x_4, x_5, x_6, x_7, x_9\}$ , unreasonable =  $\{x_8\}$ ;  
 $C$  : reasonable =  $\{x_1, x_2, x_3\}$ , ordinary =  $\{x_4, x_5, x_6, x_8, x_9\}$ , unreasonable =  $\{x_7\}$ ;  
 $D$  : reasonable =  $\{x_1, x_2, x_3\}$ , ordinary =  $\{x_4, x_5, x_6\}$ , unreasonable =  $\{x_7, x_8, x_9\}$ .

For attribute “surrounding”:

$A$  : quiet =  $\{x_1, x_2\}$ , a little noisy =  $\{x_3, x_6\}$ , noisy =  $\{x_4, x_5, x_7\}$ , quite noisy =  $\{x_8, x_9\}$ ;  
 $B$  : quiet =  $\{x_1, x_5\}$ , a little noisy =  $\{x_2, x_3\}$ , noisy =  $\{x_4, x_7, x_8\}$ , quite noisy =  $\{x_6, x_9\}$ ;  
 $C$  : quiet =  $\{x_1\}$ , a little noisy =  $\{x_2, x_3\}$ , noisy =  $\{x_4, x_7, x_8\}$ , quite noisy =  $\{x_5, x_6, x_9\}$ ;  
 $D$  : quiet =  $\{x_1, x_2, x_4\}$ , a little noisy =  $\{x_3, x_5\}$ , noisy =  $\{x_7, x_8\}$ , quite noisy =  $\{x_6, x_9\}$ .

Assume the evaluation of every specialist is of the same importance. If we want to combine these evaluations together without losing information, we should union the evaluations given by every specialist for every attribute value. Then for every attribute we get a cover instead of a partition, which embodies a kind of uncertainty.

For attribute “price” we get

$C_1 = \{\{x_1, x_2, x_4, x_5, x_7, x_8\}, \{x_2, x_5, x_8\}, \{x_2, x_3, x_5, x_6, x_8, x_9\}\}$ .

For attribute “colour” we get

$C_2 = \{\{x_1, x_2, x_3, x_4, x_5, x_6\}, \{x_4, x_5, x_6, x_7, x_8, x_9\}\}$ .

For attribute “structure” we get

$C_3 = \{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6, x_7, x_8, x_9\}, \{x_7, x_8, x_9\}\}$ .

For attribute “surrounding” we get

$C_4 = \{\{x_1, x_2, x_4, x_5\}, \{x_2, x_3, x_5, x_6\}, \{x_4, x_5, x_7, x_8\}, \{x_5, x_6, x_8, x_9\}\}$ .

Let  $\Delta = \{C_i : i = 1, \dots, 4\}$ , then

$$\Delta_1 = \{x_1, x_2\}, \Delta_2 = \{x_2\}, \Delta_3 = \{x_2, x_3\}, \Delta_4 = \{x_4, x_5\}, \Delta_5 = \{x_5\}, \Delta_6 = \{x_5, x_6\},$$

$\Delta_7 = \{x_7, x_8\}, \Delta_8 = \{x_8\}, \Delta_9 = \{x_8, x_9\}$ , here  $\Delta_i$  means  $\Delta_{x_i}$  for short, the discernibility matrix of  $(U, \Delta)$  is presented as follows:

$$\begin{pmatrix} \{C_1, C_4\} & & & & & & & & \\ \{C_1, C_4\} & \{C_1, C_4\} & & & & & & & \\ \{C_3\} & \{C_1 \wedge C_2, C_3, C_4\} & \{C_1, C_3, C_4\} & & & & & & \\ \{C_3\} & \{C_3\} & \{C_3\} & \{C_1, C_4\} & & & & & \\ \{C_1, C_3, C_4\} & \{C_1 \wedge C_2, C_3, C_4\} & \{C_3\} & \{C_1, C_4\} & \{C_1, C_4\} & & & & \\ \{C_2, C_3, C_4\} & \{C_2, C_3, C_4\} & \{C_1, C_2, C_3, C_4\} & \{C_2 \wedge C_3, C_3 \wedge C_4\} & \{C_1 \wedge C_3, C_2 \wedge C_3, C_3 \wedge C_4\} & \{C_1, C_2 \wedge C_3, C_4\} & & & \\ \{C_2, C_3, C_4\} & \{C_2, C_3, C_4\} & \{C_2, C_3, C_4\} & \{C_1 \wedge C_3, C_2 \wedge C_3, C_4\} & \{C_2 \wedge C_3, C_3 \wedge C_4\} & \{C_1 \wedge C_3, C_2 \wedge C_3, C_4\} & \{C_1, C_4\} & & \\ \{C_1, C_2, C_3, C_4\} & \{C_2, C_3, C_4\} & \{C_2, C_3, C_4\} & \{C_1, C_2 \wedge C_3, C_4\} & \{C_1 \wedge C_3, C_2 \wedge C_3, C_3 \wedge C_4\} & \{C_2 \wedge C_3, C_3 \wedge C_4\} & \{C_1, C_4\} & \{C_1, C_4\} \end{pmatrix}$$

and

$$\begin{aligned} f(U, \Delta)(\overline{C_1}, \dots, \overline{C_4}) &= \wedge \{\vee(c_{ij}) : 1 \leq j < i \leq 9, c_{ij} \neq \phi\} \\ &= C_3 \wedge (C_1 \vee C_4) \wedge (C_1 \vee C_3 \vee C_4) \wedge (C_2 \vee C_3 \vee C_4) \wedge (C_1 \vee C_2 \vee C_3 \vee C_4) \\ &\quad \wedge ((C_1 \wedge C_2) \vee C_3 \vee C_4) \wedge ((C_2 \wedge C_3) \vee (C_3 \wedge C_4)) \wedge ((C_1 \wedge C_2) \vee (C_2 \wedge C_3) \vee C_4) \\ &\quad \wedge (C_1 \vee (C_2 \wedge C_3) \vee C_4) \wedge ((C_1 \wedge C_3) \vee (C_2 \wedge C_3) \vee (C_3 \wedge C_4)) \\ &= C_3 \wedge (C_1 \vee C_4) \wedge ((C_2 \wedge C_3) \vee (C_3 \wedge C_4)) \wedge ((C_1 \wedge C_2) \vee (C_2 \wedge C_3) \vee C_4) \\ &= C_3 \wedge (C_1 \vee C_4) \wedge (C_3 \wedge (C_2 \vee C_4)) \wedge ((C_2 \wedge (C_1 \vee C_3)) \vee C_4) \\ &= C_3 \wedge (C_1 \vee C_4) \wedge (C_2 \vee C_4) \wedge ((C_2 \vee C_4) \wedge (C_1 \vee C_3 \vee C_4)) \\ &= C_3 \wedge (C_1 \vee C_4) \wedge (C_2 \vee C_4) \wedge (C_1 \vee C_3 \vee C_4) \\ &= C_3 \wedge (C_1 \vee C_4) \wedge (C_2 \vee C_4) \\ &= (C_3 \wedge C_4) \vee (C_1 \wedge C_2 \wedge C_3) \end{aligned}$$

so  $\text{RED}(\Delta) = \{\{C_3, C_4\}, \{C_1, C_2, C_3\}\}$ ,  $\text{Core}(\Delta) = \{C_3\}$ .

If these nine houses are the training samples, then we have two different kinds of evaluation references for other input samples: {structure; surrounding}, {price; colour; structure}, and it is clear that the attribute “structure” is the key attribute for the evaluation of houses.

## 5. Conclusion

In this paper we first present the applied background for covering generalized rough sets and then discuss the upper approximation of covering generalized rough sets. Our presented definition for upper approximation is more reasonable than the ones in [10,11]. We have also studied the reduct of covering generalized rough sets and propose the discernibility matrix method to compute all the reducts. This method is really the generalization of the method in [13] for traditional rough sets. We have set up the foundation of covering generalized rough set theory based on these arguments. Our future work will concentrate on the relative reducts of decision systems with respect to covering generalized rough sets.

## Acknowledgement

This work is supported by the Hong Kong RGC CERG research grants: PolyU 5193/04E (B-Q826) and PolyU 5273/05E (B-Q943). The second author is supported by the foundation of North China electric power university.

## References

- [1] Z. Pawlak, Rough Sets, International Journal of Computer and Information Science 11 (5) (1982) 341–356.
- [2] I. Jagielska, C. Matthews, T. Whitfort, An investigation into the application of neural networks, fuzzy logic, genetic algorithms, and rough sets to automated knowledge acquisition for classification problems, Neurocomputing 24 (1999) 37–54.
- [3] M. Kryszkiewicz, Rough set approach to incomplete information systems, Information Sciences 112 (1998) 39–49.
- [4] Z. Pawlak, Rough sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Boston, 1991.

- [5] S. Tsumoto, Automated extraction of medical expert system rules from clinical databases based on rough set theory, *Information Sciences* 112 (1998) 67–84.
- [6] A. Skowron, L. Polkowski, *Rough sets in knowledge discovery*, vols. 1–2, Springer, Berlin, 1998.
- [7] R. Slowinski (Ed.), *Intelligent decision support: Handbook of applications and advances of the rough sets theory*, Kluwer Academic Publishers, Boston, 1992.
- [8] W.P. Ziarko (Ed.), *Rough sets, fuzzy sets and knowledge discovery*, Workshop in Computing, Springer, London, 1994.
- [9] R. Slowinski, D. Vanderpooten, A generalized definition of rough approximations based on similarity, *IEEE Trans. Data Knowledge Eng.* 2 (2000) 331–336.
- [10] Z. Bonikowski, E. Bryniarski, U. Wybraniec, Extensions and intentions in the rough set theory, *Information Science* 107 (1998) 149–167.
- [11] J.N. Mordeson, Rough set theory applied to (fuzzy)ideal theory, *Fuzzy Sets and Systems* 121 (2001) 315–324.
- [12] William Zhu, Fei-Yue Wang, Reduct and axiomization of covering generalized rough sets, *Information Science* 152 (2003) 217–230.
- [13] A. Skowron, C. Rauszer, The discernibility matrices and functions in information systems, *Intelligent Decision support*, in: R. Slowinski (Ed.), *Handbook of applications and advances of the rough sets theory*, Kluwer Academic Publishers, 1992.
- [14] W. Zakowski, Approximations in the space  $(U, II)$ , *Demonstratio Mathematica* 16 (1983) 761–769.
- [15] Z. Bonikowski, Algebra structures of rough sets, in: W. Ziarko (Ed.), *Rough Sets, Fuzzy Sets and Knowledge Discovery*, Springer Verlag, London, 1994, pp. 243–247.
- [16] E. Bryniarski, A calculus of rough sets of the first order, *Bulletin of the Polish Academy of Science* 16 (1989) 71–77.
- [17] J.A. Pomykala, Approximation operations in approximation space, *Bulletin of the Polish Academy of Science* 9–10 (1987) 653–662.
- [18] Richard Jensen, Qiang Shen, Fuzzy-rough attributes reduction with application to web categorization, *Fuzzy Sets and Systems* 141 (2004) 469–485.
- [19] M. Kryszkiewicz, Rule in incomplete information systems, *Information Sciences* 113 (1999) 271–292.
- [20] J.M. Fernandez Salido, S. Murakami, Rough set analysis of a general type of fuzzy data using transitive aggregations of fuzzy similarity relations, *Fuzzy Sets and Systems* 139 (2003) 635–660.
- [21] Chen Degang, Zhang Wenxiu, D.S. Yeung, E.C.C. Tsang, Rough approximations on a complete completely distributive lattice with applications to generalized rough sets, *Information Sciences* 176 (2006) 1829–1848.
- [22] Rajen B. Bhatt, M. Gopal, On fuzzy rough sets approach to feature selection, *Pattern recognition Letters* 26 (11) (2005) 1632–1640.